Solving PhaseLift by low-rank Riemannian optimization methods for complex semidefinite constraints

Wen Huang
Université Catholique de Louvain

Kyle A. Gallivan
Florida State University

Xiangxiong Zhang
Purdue University

March 2015

The Phase Retrieval Problem

- The Phase Retrieval problem concerns recovering a signal given the modulus of its linear transform;
- It is important in many applications, e.g.,
  - X-ray crystallography imaging [Har93];
  - Diffraction imaging [BDP+07];
  - Optics [Wal63];
  - Microscopy [MISE08];
- The Fourier transform is considered;

Problem Statement

- Recover the signal $x : \mathbb{R}^s \to \mathbb{C}$ from intensity measurements of its Fourier transform, $|\tilde{x}(u)| = \left| \int_{\mathbb{R}^s} x(t) \exp(-2\pi u \cdot t\sqrt{-1}) dt \right|$;
- Discrete form

$$\text{find } \mathbf{x} \in \mathbb{C}^{n_1 \times n_2 \times \ldots \times n_s}, \text{ s. t. } |Ax| = b,$$

where $x = \mathrm{vec}(\mathbf{x}) \in \mathbb{R}^n$, $n = n_1 n_2 \cdots n_s$ and $A \in \mathbb{C}^{m \times n}$ defines the Discrete Fourier transform;

Difficulties and Oversampling

- Solution of the discrete form may be not unique.
- Oversampling in the Fourier domain is a standard method to obtain a unique solution.
    - No benefit for most 1D signals, see e.g., [San85].
    - Give a unique solution for multiple dimensional problems for constrained signals, see e.g. [BS79, Hay82, San85].
    - Algorithms based on alternating projection are used.

Other Frameworks

- PhaseLift [CESV13] and PhaseCut [WDM13]: Combining using multiple structured illuminations or masks with convex programming;
- A unique rank one solution up to a global phase factor [CESV13, CL13, CSV13, WDM13];
- Stability [CL13, CSV13, WDM13];
- Convex programming solvers, e.g., SDPT3 [TTT99] or TFOCS [BCG11];
- The PhaseLift framework is considered in this presentation.

PhaseLift: Using Illumination Fields

- The known illumination fields on the discrete signal domain $\mathbf{w}_r \in \mathbb{C}^{n_1 \times n_2 \times \dots n_s}$, $r = 1, \dots \ell$.
- Let w_r denote $\mathrm{vec}(\mathbf{w}_r)$. One illumination field gives an equation

$$|A\,\mathrm{Diag}(w_r)x| = b_r$$

  where $\mathrm{Diag}(w_r)$ denotes an n-by-n diagonal matrix the diagonal entries of which are w_r.
- ℓ fields yields

$$\left| \begin{pmatrix} A & & & \\ & A & & \\ & & \ddots & \\ & & & A \end{pmatrix} \begin{pmatrix} \mathrm{Diag}(w_1) \\ \mathrm{Diag}(w_2) \\ \vdots \\ \mathrm{Diag}(w_\ell) \end{pmatrix} x \right| = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_\ell \end{pmatrix} := b$$

PhaseLift: Using Illumination Fields

- Therefore, the linear operator A for PhaseLift using the Fourier transform, denote Z, is

$$Z = \begin{pmatrix} A & & & \\ & A & & \\ & & \ddots & \\ & & & A \end{pmatrix} \begin{pmatrix} \mathrm{Diag}(w_1) \\ \mathrm{Diag}(w_2) \\ \vdots \\ \mathrm{Diag}(w_\ell) \end{pmatrix}$$

- PhaseLift problem:

$$\text{find } x \in \mathbb{C}^n,$$
$$\text{s. t. } |Zx| = b, \text{ or equivalently, } \mathrm{diag}(Zxx^* Z^*) = b^2,$$

where $*$ denotes the conjugate transpose operator.

PhaseLift: Lifting to Convex Problem

- Let $X \in \mathbb{C}^{n \times n}$ denote xx^*. The Phase Retrieval problem becomes

$$\text{find } X, \quad \text{s. t. } \operatorname{diag}(ZXZ^*) = b^2, X \geq 0 \text{ and } \operatorname{rank}(X) = 1,$$

  or equivalently

$$\min \operatorname{rank}(X), \quad \text{s. t. } \operatorname{diag}(ZXZ^*) = b^2, \text{ and } X \geq 0,$$

  where $X \geq 0$ denotes X is Hermitian positive semidefinite.
- Convex programming

$$\min \operatorname{trace}(X), \quad \text{s. t. } \operatorname{diag}(ZXZ^*) = b^2, \text{ and } X \geq 0.$$

PhaseLift: Noise Measurements

- Measurements with noise, $b^2 \in \mathbb{R}^m$, are sampled from a probability distribution $p(\cdot;\mu)$, where $\mu = \mathrm{diag}(ZXZ^*)$.
- Minimize the negative log-likelihood function

$$\min_x -\log(p(b;\mu))$$
$$\text{such that } \mu = \mathrm{diag}(Zxx^*Z^*),$$

- Similarly, an alternate problem can be used:

$$\min -\log(p(b;\mu)) + \kappa\,\mathrm{trace}(X), \quad \text{s. t. } \mathrm{diag}(ZXZ^*) = b^2, \text{ and } X \geq 0.$$

  where κ is a positive constant.

- If the likelihood is log-concave, then it is a convex problem, e.g., for Poisson or Gaussian distributions.

PhaseLift: Nonconvex Approach

- The complexity can be too high in convex approach.
- The alternate problems are

$$\text{noiseless:} \quad \min_{X \geq 0} \|b^2 - \mathrm{diag}(ZXZ^*)\|_2^2 + \kappa \, \mathrm{trace}(X),$$

$$\text{noise:} \quad \min_{X \geq 0} - \log(p(b; \mathrm{diag}(ZXZ^*))) + \kappa \, \mathrm{trace}(X)$$

  where κ is a positive constant;
- They are used in [CESV13] and reweighting is used to promote low-rank solutions;
- This motivates us to consider the optimization problem

$$\min_{X \geq 0} H(X)$$

  and the desired minimizer is low rank. In particular for the PhaseLift problem, the rank of desired minimizer is 1.

Optimization on Hermitian Positive Semidefinite Matrices

- Suppose the rank of desired minimizer r^* is at most p.
- The domain $\{X \in \mathbb{C}^{n\times n} | X \geq 0\}$ can be replaced by $\mathcal{D}_p$, where $\mathcal{D}_p = \{X \in \mathbb{C}^{n\times n} | X \geq 0, \mathrm{rank}(X) \leq p\}$.
- An alternate cost function can be used:

$$F_p : \mathbb{C}^{n\times p} \to \mathbb{R} : Y_p \mapsto H(Y_p Y_p^*).$$

- Note that for the PhaseLift problem, choosing $p = 1$ is equivalent not to do the Lifting step. Choosing $p > 1$ yields computational and theoretical benefits.
- This idea is not new and has been discussed in [BM03] and [JBAS10] for real positive semidefinite matrix constraints.

First Order Optimality Condition

Theorem

If $Y_p^* \in \mathbb{C}^{n \times p}$ is a rank deficient minimizer of F_p, then $Y_p Y_p^*$ is a stationary point of H.
In addition, if H is a convex cost function, $Y_p Y_p^*$ is a global minimizer of H.

- The real version of the optimality condition is given in [JBAS10].

Optimization Framework

- Equivalence: if $Y_p Y_p^* = \tilde{Y}_p \tilde{Y}_p^*$, then $F_p(Y_p) = F_p(\tilde{Y}_p)$;
- Quotient manifolds are used to remove the equivalence:
  - Equivalent class of $Y_r \in \mathbb{C}_*^{n \times r}$ is $[Y_r] = \{Y_r O_r | O_r \in \mathcal{O}_r\}$, where $1 \leq r \leq p$, $\mathbb{C}_*^{n \times r}$ denotes the n-by-r complex noncompact Stiefel manifold and $\mathcal{O}_r$ denote the r-by-r complex rotation group;
  - A fixed rank quotient manifold $\mathbb{C}_*^{n \times r}/\mathcal{O}_r = \{[Y_r] | Y_r \in \mathbb{C}_*^{n \times r}\}$, $1 \leq r \leq p$;
- Function on a fixed rank manifold is

$$f_r : \mathbb{C}_*^{n \times r}/\mathcal{O}_r \to \mathbb{R} : [Y_r] \mapsto F_r(Y_r) = H(Y_r Y_r^*);$$

- Optimize the cost function f_r and update r if necessary;
- A similar approach is used in [JBAS10] for real problems;

Update Rank Strategy

- Most of work is to choose a upper bound k for the rank and optimize over $\mathbb{C}^{n\times k}$ or $\mathbb{R}^{n\times k}$.
- Increasing rank by a constant [JBAS10, UV14]
  - Descent
  - Globally converge
- Dynamically search for a suitable rank [ZHG$^+$15]
  - Not descent
  - Globally converge

Update Rank Strategy

- Rank reduce is used for the problem in PhaseLift;
- The rank is reduced if the singular values of an iterate have notable bias:
  - Suppose r is the rank of current iterate and $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_r$ are its singular values;
  - Given a threshold $\delta \in (0, 1)$, the next rank is q if $\sigma_q > \delta\tilde{\sigma}$ and $\sigma_{q+1} \leq \delta\tilde{\sigma}$, where $\tilde{\sigma} = \|\operatorname{Diag}(\sigma_1, \ldots, \sigma_r)\|_F / \sqrt{r}$;
- The next iterate is given by truncating relative small singular values;

Riemannian Optimization

Riemannian optimization algorithms are used to optimize the problem on the fixed rank manifold $\mathbb{C}_*^{n\times r}/\mathcal{O}_r$.

Line search Newton (RNewton)	[AMS08]
Trust region Newton (RTR-Newton)	[Bak08]
BFGS (RBFGS)	[RW12, HGA14]
Limited memory version of BFGS (LRBFGS)	[HGA14]
Trust region symmetric rank one update method (RTR-SR1)	[HAG14]
Limited memory version of RTR-SR1 (LRTR-SR1)	[HAG14]
Riemannian conjugate gradient method (RCG)	[AMS08]

Algorithm 1

Set initial rank $r = p$;

for $k = 0, 1, 2, \ldots$ **do**

Apply Riemannian method for cost function f_r over $\mathbb{C}_*^{n \times r} / \mathcal{O}_r$ with initial point $[Y_r^{(k)}]$ until i-th iterate $[W^{(i)}]$ satisfying $\| \operatorname{grad} f_r([W^{(i)}]) \| < \epsilon$ or the requirement of reducing rank with threshold δ;

if $\| \operatorname{grad} f_r([W^{(i)}]) \| < \epsilon$ **then**

Find a minimizer $[W] = [W^{(i)}]$ over $\mathbb{C}_*^{n \times p} / \mathcal{O}_p$ and return;

else {iterate in the Riemannian optimization method meets the requirements of reducing rank}

Reduce the rank to $q < r$ based on truncation with threshold δ and obtain an output $\hat{W} \in \mathbb{C}^{n \times q}$;

$r \leftarrow q$ and set $[Y_r^{(k+1)}] = [\hat{W}]$;

end if

end for

Numerical Experiments

- Artificial Data sets
  - The entries of true solution x_* and the masks $w_i, i = 1, \ldots, l$ are drawn from the standard normal distribution;
  - x_* is further normalized by $\|x_*\|_2$;
  - $w_i, i = 1, \ldots, l$ is further normalized by $\sqrt{n}$;
  - The measurements b^2 is set to be $\mathrm{diag}(Z x_* x_*^* Z^*) + \epsilon$, where the entries of $\epsilon \in \mathbb{R}^m$ are drawn from the normal distribution with mean 0 and variance τ.

Cost function and Complexities

- The cost function in this case is

$$f_r([Y_r]) = \frac{\|b^2 - \mathrm{diag}(ZY_rY_r^*Z^*)\|_2^2}{\|b^2\|_2^2} + \kappa\,\mathrm{trace}(Y_rY_r^*);$$

- The closed forms of gradient and action of Hessian are known [HGZ14];
- Their complexities respectively are
  - Function evaluation: $O(\ell pns\max_i(log(n_i)))$;
  - Gradient evaluation: $O(\ell pns\max_i(log(n_i))) + O(np^2) + O(p^3)$;
  - Action of Hessian: $O(\ell pns\max_i(\log(n_i)) + O(np^2) + O(p^3)$;
- If $p << n$ (it is true in practice), then all these complexities are dominated by $O(\ell pns\max_i(\log(n_i))$;

Default Setting

- All tests are performed in Matlab R2014a on a 64 bit Ubuntu system with 3.6 GHz CPU (Intel (R) Core (TM) i7-4790).
- The stopping criterion requires the norm of gradient to less than 10^{-6};
- The number of masks ℓ is 6;
- The coefficient κ in the cost function is 0;
- Threshold $\delta = 0.9$ for rank reduction;

Representative Riemannian Algorithms

- RNewton, RTR-Newton, LRBFGS, LRTR-SR1 and RCG;
- Noiseless measurements, i.e., $\tau = 0$;
- Initial rank $p_0 = 8$;
- Average of 10 random runs;

Representative Riemannian Algorithms

Table : $n = n_1 n_2$. The subscript ν indicates a scale of 10^{ν}.

(n_1, n_2)		(32, 32)	(32, 64)	(64, 64)	(64, 128)	(128, 128)	(128, 256)	(256, 256)
RNewton	nf	3.24_1	3.36_1	3.58_1	4.12_1	4.58_1	5.74_1	5.88_1
	ng	2.34_1	2.58_1	2.92_1	3.26_1	3.8_1	4.56_1	5.14_1
	nH	4.56_2	3.39_2	3.86_2	4.00_2	4.21_2	4.88_2	5.73_2
	f_f	1.25_{-13}	2.17_{-13}	1.37_{-14}	4.83_{-13}	1.59_{-14}	2.32_{-12}	3.79_{-13}
	t	2.54	3.25	9.28	1.72_1	3.02_1	7.57_1	1.98_2
RTR-Newton	nf	2.82_1	2.66_1	2.84_1	3.06_1	3.42_1	3.62_1	3.36_1
	ng	2.82_1	2.66_1	2.84_1	3.06_1	3.42_1	3.62_1	3.36_1
	nH	3.00_2	6.13_2	4.36_2	484	5.27_2	5.33_2	5.77_2
	f_f	2.11_{-14}	2.86_{-13}	4.53_{-13}	4.69_{-13}	3.03_{-13}	9.74_{-14}	3.45_{-14}
	t	1.83	5.45	8.23	1.92_1	3.42_1	6.07_1	1.42_2
LRBFGS	nf	9.78_1	1.06_2	1.20_2	133	1.45_2	1.84_2	191
	ng	9.6_1	1.04_2	1.16_2	1.29_2	1.40_2	1.77_2	1.87_2
	f_f	6.40_{-12}	6.82_{-12}	7.61_{-12}	1.04_{-11}	1.58_{-11}	2.24_{-11}	3.55_{-11}
	t	6.00_{-1}	1.03	1.90	3.37	6.86	1.59_1	3.27_1
LRTR-SR1	nf	1.45_2	1.44_2	1.56_2	1.71_2	1.88_2	2.29_2	2.43_2
	ng	1.45_2	1.44_2	1.56_2	1.71_2	1.88_2	2.29_2	2.43_2
	f_f	1.24_{-11}	1.08_{-11}	1.50_{-11}	3.67_{-11}	3.10_{-11}	4.82_{-11}	2.04_{-10}
	t	9.97_{-1}	1.64	3.16	6.20	1.22_1	3.14_1	6.81_1
RCG	nf	2.66_2	2.59_2	2.77_2	2.89_2	3.11_2	3.45_2	3.73_2
	ng	2.55_2	250	266	2.80_2	3.02_2	3.36_2	3.65_2
	f_f	3.11_{-12}	3.47_{-12}	5.42_{-12}	7.94_{-12}	1.16_{-11}	1.60_{-11}	2.53_{-11}
	t	1.18	2.00	3.74	7.18	1.47_1	3.63_1	9.54_1

LRBFGS is chosen to be the representative Riemannian algorithm.

Compare with a Convex Programming Solver

- Compare with convex programming
  - FISTA [BT09] in Matlab library TFOCS [BCG11];
  - X can be too large to be handled by the solver;
  - A low rank version of FISTA is used, denoted by LR-FISTA;
  - The approach is used in [CESV13, CSV13];
  - Works in practice but no theoretical results.

Compare with a Convex Programming Solver

- $n_1 = n_2 = 64$; $n = n_1 n_2 = 4096$;
- LR-FISTA stops if $\frac{\|X^{(i)} - X^{(i-1)}\|_F}{\|X^{(i)}\|_F} < 10^{-6}$ or *iter* > 2000;
- Noise and noiseless problems are tested;
- For noise measurements:
  - $\tau = 10^{-4}$, i.e., the signal-to-noise ration (SNR) $10 \log_{10} \left(\frac{\|b^2\|_2^2}{\|b^2 - \hat{b}^2\|_2^2} \right)$ is 31.05 dB, where $b^2 = \mathrm{diag}(Z x_* x_*^* Z^*)$ and $\hat{b}$ is the noise measurements;
  - Multiple κ are used;

Noiseless Measurements

Table : k denotes the upper bound of the low-rank approximation in LR-FISTA. $\sharp$ represents the number of iterations reach the maximum. The relative mean-square error (RMSE) is $\min_{a:|a|=1} \|ax - x_*\|_2/\|x_*\|_2$.

noiseless	Algorithm 1	LR-FISTA (k)				
		1	2	4	8	16
iter	124	1022	377	601	1554	$2000^\sharp$
nf	129	2212	804	1278	3360	4322
ng	124	1106	402	639	1680	2161
f_f	4.62_{-12}	8.18_{-12}	4.50_{-11}	4.64_{-12}	1.54_{-11}	1.27_{-9}
RMSE	6.34_{-6}	1.01_{-5}	1.74_{-5}	1.46_{-5}	1.10_{-4}	2.56_{-3}
t	2.12	1.27_2	5.25_1	9.35_1	3.48_2	6.86_2

- Algorithm 1 is faster and gives smaller RMSE.

Noise Measurements

Table : k denotes the upper bound of the low-rank approximation in LR-FISTA. RMSE denotes $\min_{a:|a|=1} \|ax - x_*\|_2 / \|x_*\|_2$. $\sharp$ represents the number of iterations reach the maximum.

noise	κ	Algorithm 1	LR-FISTA (k)				
			1	2	4	8	16
iter	10^{-2}	84	409	$2000^{\sharp}$	$2000^{\sharp}$	$2000^{\sharp}$	$2000^{\sharp}$
	10^{-4}	122	978	$2000^{\sharp}$	$2000^{\sharp}$	$2000^{\sharp}$	$2000^{\sharp}$
nf	10^{-2}	86	886	4280	4284	4290	4280
	10^{-4}	129	2116	4296	4316	4300	4318
ng	10^{-2}	84	526	3468	3376	3242	3371
	10^{-4}	122	1105	2148	2158	2150	2159
f_f	10^{-2}	1.63_{-1}	1.63_{-1}	1.77_{-1}	2.24_{-1}	2.75_{-1}	3.04_{-1}
	10^{-4}	1.80_{-3}	1.80_{-3}	1.81_{-3}	2.19_{-3}	4.55_{-3}	7.01_{-3}
RMSE	10^{-2}	1.80_{-1}	1.80_{-1}	2.64_{-1}	3.60_{-1}	4.19_{-1}	4.45_{-1}
	10^{-4}	2.63_{-3}	2.63_{-3}	6.46_{-3}	2.17_{-2}	4.98_{-2}	6.57_{-2}
t	10^{-2}	1.59	5.17_{1}	3.79_{2}	4.48_{2}	5.73_{2}	9.45_{2}
	10^{-4}	2.06	1.21_{2}	3.05_{2}	3.17_{2}	4.80_{2}	7.85_{2}

Noise Measurements (Continue)

Table : k denotes the upper bound of the low-rank approximation in LR-FISTA. RMSE denotes $\min_{a:|a|=1} \|ax - x_*\|_2 / \|x_*\|_2$. $\sharp$ represents the number of iterations reach the maximum.

noise	κ	Algorithm 1	LR-FISTA (k) 1	2	4	8	16
iter	10^{-6}	128	1027	$2000^\sharp$	$2000^\sharp$	$2000^\sharp$	$2000^\sharp$
	0	138	1070	$2000^\sharp$	$2000^\sharp$	$2000^\sharp$	$2000^\sharp$
nf	10^{-6}	132	2210	4266	4312	4336	4316
	0	143	2306	4308	4322	4314	4320
ng	10^{-6}	128	1105	2712	2156	2168	2158
	0	138	1153	2154	2161	2157	2160
f_f	10^{-6}	1.84_{-5}	1.84_{-5}	1.91_{-5}	2.35_{-5}	3.55_{-5}	7.62_{-5}
	0	4.08_{-7}	4.08_{-7}	1.16_{-6}	6.27_{-6}	2.51_{-5}	8.89_{-5}
RMSE	10^{-6}	6.72_{-4}	6.72_{-4}	1.09_{-3}	2.10_{-3}	3.53_{-3}	6.27_{-3}
	0	6.70_{-4}	6.70_{-4}	1.09_{-3}	2.18_{-3}	4.01_{-3}	7.29_{-3}
t	10^{-6}	2.13	1.27_2	2.75_2	3.01_2	4.64_2	7.04_2
	0	2.20	1.34_2	2.63_2	2.98_2	4.32_2	6.91_2

The Gold Ball Data

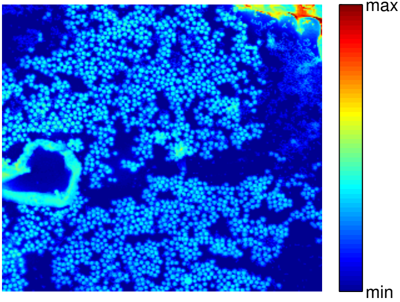

Figure : Image of the absolute value of the 256-by-256 complex-valued image. $n = 65536$. The pixel values correspond to the complex transmission coefficients of a collection of gold balls embedded in a medium.

Thank Stefano Marchesini at Lawrence Berkeley Notional Laboratory for providing the gold balls data set and granting permission to use it.

The Gold Ball Data

A set of binary masks contains a mask that is all 1 (which yields the original image) and several other masks comprising elements that are 0 or 1 with equal probability.

Table : RMSE and computational time (second) results with varying number and types of masks are shown in format RMSE/TIME. $\sharp$ represents the computational time reaching 1 hour, i.e., 3.6_3 seconds.

	Algorithm 1			LR-FISTA		
SNR (dB)	20	40	inf	20	40	inf
6 Gaussian	$8.32_{-3}/4.30_1$	$8.32_{-5}/4.50_1$	$3.12_{-6}/4.19_1$	$8.32_{-3}/\sharp$	$3.12_{-4}/\sharp$	$3.12_{-4}/\sharp$
6 binary	$7.23_{-1}/7.90_2$	$1.29_{-1}/4.24_2$	$1.09_{-1}/4.42_2$	$8.24_{-1}/\sharp$	$4.98_{-1}/\sharp$	$4.98_{-1}/\sharp$
32 binary	$2.21_{-1}/6.84_2$	$3.02_{-3}/7.36_2$	$2.57_{-3}/6.54_2$	$6.07_{-1}/\sharp$	$5.82_{-1}/\sharp$	$5.78_{-1}/\sharp$

The Gold Ball Data

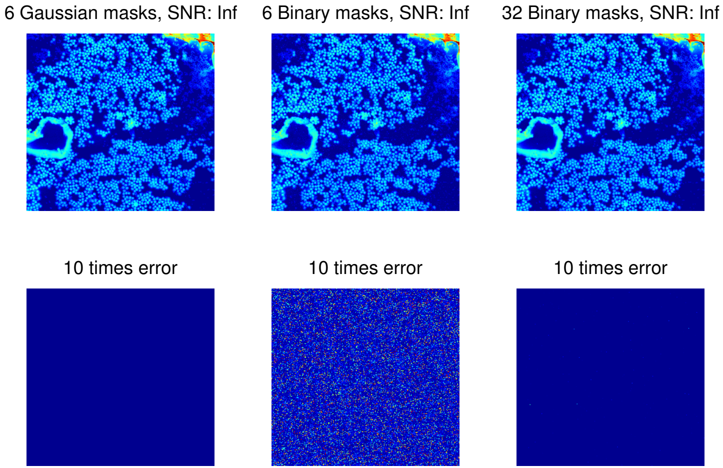

Figure : Reconstructions via PhaseLift with varying number and types of masks. For the same number and types of masks, the reconstructions of

The Gold Ball Data

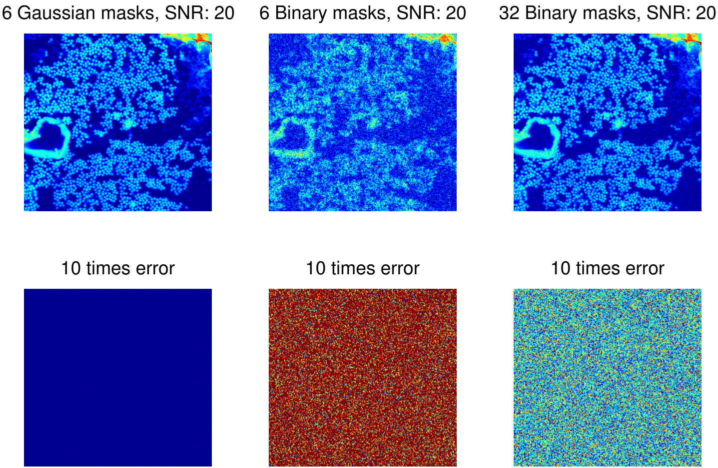

Figure : Reconstructions via PhaseLift with varying number and types of masks. For the same number and types of masks, the reconstructions of noisy

Conclusion

- A low-rank problem is proposed to replace optimization problems on Hermitian positive semidefinite matrices;
- The first order optimality condition is given;
- For the PhaseLift problem, an algorithm based on a rank reduce strategy and a state-of-the-art Riemannian algorithm is suggested;
- Experiments of noise, noiseless, Gaussian masks and binary masks are tested and show that the new algorithm is more efficient and effective than the LR-FISTA algorithm.

References I

P.-A. Absil, R. Mahony, and R. Sepulchre.
Optimization algorithms on matrix manifolds.
Princeton University Press, Princeton, NJ, 2008.

C. G. Baker.
Riemannian manifold trust-region methods with applications to eigenproblems.
PhD thesis, Florida State University, Department of Computational Science, 2008.

S. Becker, E. J. Cand, and M. Grant.
Templates for convex cone problems with applications to sparse signal recovery.
Mathematical Programming Computation, 3:165–218, 2011.

References II

O. Bunk, A. Diaz, F. Pfeiffer, C. David, B. Schmitt, D. K. Satapathy, and J. F. van der Veen.
Diffractive imaging for periodic samples: retrieving one-dimensional concentration profiles across microfluidic channels.
Acta crystallographica. Section A, Foundations of crystallography, 63(Pt 4):306–314, July 2007.
doi:10.1107/S0108767307021903.

S. Burer and R. D. C. Monteiro.
A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization.
Mathematical Programming, 95(2):329–357, February 2003.
doi:10.1007/s10107-002-0352-8.

Y. M. Bruck and L. G. Sodin.
On the ambiguity of the image recontruction problem.
Optics Communications, 30(3):304–308, 1979.

A. Beck and M. Teboulle.
A fast iterative shrinkage-thresholding algorithm for linear inverse problems.
SIAM Journal on Imaging Sciences, 2(1):183–202, January 2009.
doi:10.1137/080716542.

E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski.
Phase retrieval via matrix completion.
SIAM Journal on Imaging Sciences, 6(1):199–225, 2013.
arXiv:1109.0573v2.

E. J. Candès and X. Li.
Solving quadratic equations via phaselift when there are about as many equations as unknowns.
Foundations of Computational Mathematics, June 2013.
doi:10.1007/s10208-013-9162-z.

References IV

E. J. Candès, T. Strohmer, and V. Voroninski.
PhaseLift : Exact and stable signal recovery from magnitude measurements via convex programming.
Communications on Pure and Applied Mathematics, 66(8):1241–1274, 2013.

W. Huang, P.-A. Absil, and K. A. Gallivan.
A Riemannian symmetric rank-one trust-region method.
Mathematical Programming, February 2014.
doi:10.1007/s10107-014-0765-1.

R. W. Harrison.
Phase problem in crystallography.
Journal of the Optical Society of America A, 10(5):1046–1055, May 1993.
doi:10.1364/JOSAA.10.001046.

References V

M. H. Hayes.
The reconstruction of a multidimensional sequence from the phase or magnitude of its fourier transform.
IEEE Transactions on Acoustics speech and signal processing, 30(2):140–154, 1982.

W. Huang, K. A. Gallivan, and P.-A. Absil.
A Broyden class of quasi-Newton methods for Riemannian optimization.
Submitted for publication, 2014.

Wen Huang, K. A. Gallivan, and X Zhang.
Solving Phaselift by low-rank Riemannian optimization methods for complex semidefinite constraints.
Submitted for publication, 2014.

M. Journée, F. Bach, P.-A. Absil, and R. Sepulchre.
Low-rank optimization on the cone of positive semidefinite matrices.
SIAM Journal on Optimization, 20(5):2327–2351, 2010.

References VI

J. Miao, T. Ishikawa, Q. Shen, and T. Earnest.
Extending X-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes.
Annual review of physical chemistry, 59:387–410, January 2008. doi:10.1146/annurev.physchem.59.032607.093642.

W. Ring and B. Wirth.
Optimization methods on Riemannian manifolds and their application to shape space.
SIAM Journal on Optimization, 22(2):596–627, January 2012. doi:10.1137/11082885X.

J. L. C. Sanz.
Mathematical considerations for the problem of Fourier transform phase retrieval from magnitude.
SIAM Journal on Applied Mathematics, 45(4):651–664, 1985.

References VII

K. C. Toh, M. J. Todd, and R. H. Tutuncu.
SDPT3 a matlab software package for semidefinite programming.
Optimization methods and software, 11:545–581, 1999.

A. Uschmajew and B. Vandereycken.
Line-search methods and rank increase on low-rank matrix varieties.
In *Proceedings of the 2014 International Symposium on Nonlinear Theory and its Applications (NOLTA2014)*, 2014.

A. Walther.
The question of phase retrieval in optics.
Optica Acta: International Journal of Optics, 10(1):41–49, January 1963.
doi:10.1080/713817747.

I. Waldspurger, A. DAspremont, and S. Mallat.
Phase recovery, maxcut and complex semidefinite programming.
Mathematical Programming, December 2013.
doi:10.1007/s10107-013-0738-9.

References VIII

G. Zhou, W. Huang, K. A. Gallivan, P. Van Dooren, and P.-A. Absil.
Rank-constrained optimization: A Riemannian manifold approach.
In *The 23th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), Accepted*, number 1, pages 1–6, 2015.